Pashto Transliteration System

BGN/PCGN 1968 System, 2006 Revision

This system was developed in 1968 for the romanization of Pashto geographic names. Dari names in Afghanistan are romanized in accordance with the Romanization System for Persian (BGN/PCGN 1958 System).

The Pashto alphabet uses a modified form of the Perso-Arabic script, and contains eight additional consonants not present in standard Arabic, as well as two additional vowel characters and an additional vowel point. Since both Pashto and Persian are official languages in Afghanistan, the same conversion values are used for all consonant characters which are common to the two languages. The eight additional Pashto consonant characters (shown below) are represented by Roman-script conversion values which suggest approximate pronunciations, and which, in six instances, bear sub-macrons.

ت څ ځ ډ ړ ښ ږ ن ي Consonants:

The two additional vowels and the one vowel point found in Pashto but not in Arabic are shown below:

Vowels: シ Vowel Point: つ

The Pashto system of vowel pointing resembles that of both Arabic and Persian but contains additional features found in neither language. As in other Perso-Arabic-script writing systems, the vowel points and some of the diacritical marks are generally omitted from both manual and machine writing and where used are rarely fully applied. Consequently, a reference source may sometimes be required to ensure correct identification of the standard spelling and proper vowels. It is also necessary to ensure elimination of dialectal and idiosyncratic variations. In the interest of clarity, full vowel pointing has been applied to the examples below.

(Note: it is recommended that a font such as Scheherezade, available from www.sil.org, which includes the entire Unicode extended Arabic subrange, be used to view this system.)

Table 1: Consonant Characters

	Pushto				Romanization	Examples and Remarks		
	Final	Medial	Initial	Independent				
1.	l			١	not romanized in word-initial position		See note 1	
2.	ب ب	÷	÷	ų.	b	بَعْلَان	Baghlān	
3.	پ	1	ڔٙ	پ	р	ۑؙۅ۫ؾٞػؽ	Pow <u>t</u> akay	
4.	ت	ت	ت	ت	t	شِرین تگاب	Shirīn Tagāb	
5.	ټ	ت	ټ	Ţ	<u>t</u>	ټڼي	<u>T</u> a <u>n</u> ĭ	
6.	ٹ	<u> </u>	ڎٞ	ث	Ē	ثابر	Šābir	

		Pushto		Romanization	Examples a	ind Remarks		
	Final	Medial	Initial	Independent				
	ج	ج	ج	ج	j	جَلَال آبَاد	Jalālābād	
8.	<u>~</u> _	~	ر ئ ا	چ	ch	چَارِيگار	Chārīkār	
9.	ۓڅ	*	څ	چ څ	ts	خُوكي	Tsawkêy	
10.	ـځ	ځ	ځ	ځ	Dz	خُدْرَان	Dzadrăn	
11.	ـخ	<u> </u>	<u> </u>	خ	Dz	جَدْران	Dzadrān	
12.	حح			ح	ļ	حَضَرَت إمَام	Ḥazarat Imām	
13.	ـخ	خ	خ	خ	kh	خُوْسْت	Khowst	
14.	7			7	d	سپين بُوْلدَک	Spīn Bowldak	
15.				Ĵ	<u>d</u>	چند و پَتان	<u>D</u> and wa Patān	
16.	خ			ذ	Z	دُوْبَى غَر	Zowbay Ghar	
17.	ـر			ر	r	گنْدَهَار	Kandahār	
18.				ړ	<u>r</u>	ائدَر	Anda <u>r</u>	
19.	۔ز			ز	Z	كُنْدُز	Kunduz	
20.	ڗ			ژ	zh	مِيرٍ أَسْلَم رُّرَ نْدَه	Mī <u>r</u> Aslam Zhrandah	
21.	۔و			و	<u>zh</u>	ڔؚؽؘۯؽڹ	<u>Zh</u> īyarayn	
22.	س	ـىد	سد	س	s	دسَلِیم کلی	Dê Salīm Kêlay	
23.	ش	شد	شد	ش	sh	مَزَارُ شَرِيف	Mazāri Sharif	
24.	ـش	ــبتــ	ىتىــ	ۺ	<u>sh</u>	كْښَتَه كَلَا	K <u>sh</u> atah Kalā	
25.	ـص	ے۔	صد	ص	ş	قَيْصَار	Qayşār	
26.	_ض	ے	ضد	ض	Z	فَيْض آبَاد	Fay <u>z</u> ābād	
27.	ط	ط	ط	ط	ţ	قطسِئ سُفلی	Quţsī-yi Suflá	
28.	ظ	ظ	ظ	ظ	Ż	ظاهِرگلا	Zāhir Kalā	
29.	ےع	2	2	ع	٤	پُلُ عَلْم	Pul-i 'Alam	
30.	ـغ	غ	غ	غ	gh	غَرْنِي	Ghaznī	
31.	ف	ف	؋	ف	f	مَزَارُ شَرِيف	Mazāri Sharīf	
32.	ـق	ق	; 9	ق	q	ققطىيئ سُفلى	Quţsī-yi Suflá	
33.	ے	5.	≥	ک	k	گنْدَهَار	Kandahār	
34.	ےگ	ےگ	گ	گ	g	گرْدېز	Gardēz	
35.	ٹ	ے	ے	ګ	g	گِيزاب	Gīzāb	
36.	上	Ţ	7	ل	I	گابَل	Kābêl	
37.	ے	؎	م	م	m	مَیْمَنَه	Maymanah	
38.	ٺ	نـ	ذ	ن	n	خَان آبَاد	Khānābād	

39.	ڹ	ڹ	ب	ڼ	<u>n</u>	ټَلْبِي	<u>T</u> a <u>n</u> ī
40.	۔و			و	w	وَاخَان	Wākhān
41.	هـ	8	ھ	٥	h	هِرَات	Hirāt
42.	ے	Ť	ڌ	ی	у	يُوسُف خبِل	Yūsuf Khēl
43.	_ي	Ť	ڌ	ي	у	پَکثیَا	Paktyā

Table 2: Vowel Characters

	Pushto	Romanization	Examples and Remarks				
1.	,	а	گابَل	Kābêl			
2.	ľ	ā	هِرَات	Hirāt			
3.	-	i	هِرَات	Hirāt			
4.	۔ ی	ī	غَرْ نِي	Ghaznī			
5.	۔ ي	ī	ڰؚيزَاب	Gīzāb			
6.	ځي	ay	شُكُورْ كُلِّي	Shakūr Kêlay			
7.	ۦێ	ey		(deprecated)			
8.	ې	ē	گرْدېز	Gardēz			
9.	وَ	aw	څُوْکۍ	Tsawkêy			
10.	َ او	āw	يَكاوْلنْگ	Yakāwlang			
11.	۔و	ew		(deprecated)			
12.	<u>-</u> ~	Ê	شلگر	Shêlgar			
13.	ۍ	êy	سَرْگانئ	Sarkānêy			
14.	ئى	êy	گوئ	kawêy, <i>you do</i> ^{4,10}			
15.	ş	U	مُرْغَاب	Murghāb			
16.	و	Ū	بَالَا بُلُوك	Bālā Bulūk			
17.	<u>ُ</u> و	ow	ڵۅ۠ڲڔ	Lowgar			
18.	<i>وی</i>	ūy	شًاه جُوى	Shāh Jūy			
19.	ایَ	āy	دَايْمِيرْ دَاد	Dāymīrdād			
20.	Ĩ	Ā	جَلال آبَاد	Jalālābād			
21.	۶	,	خِدير ٔ پَائيت	Khidīr-i Pā'īn			
22.	- -	doubling of conso- nant letter	خَانِّشِين	Khānnishīn			
23.	ئ	Á	قطىيى ئىقلى	Qutsī-yi Suflá			

			,	NUME	RALS	,			
•	١	٢	٣	٤, ۴	۵,۵	۶ ,٦	V	٨	٩
0	1	2	3	4	5	6	7	8	9

Although Perso-Arabic script is written from right to left, numerical expressions, e.g. \٩₹٨ → 1968

are written from left to right.

NOTES

- 1. Alēf (1) should be Romanized as follows:
 - a. In word-initial position, alēf should not be Romanized.
 - b. In word-initial position, alēf with maddah () should be Romanized ā. (See the vowel table.)
 - c. In word-medial and word-final position, alef should be Romanized a. (See the vowel table.)
 - d. In word-medial and word-final position in words of Arabic origin, alēf may serve as the bearer of hamzah, e.g. رأس ra's (See note 6.)
- 2. The characters tsē (خُ) and dzē (خُ) may be Romanized tŝ and dæ when for special reasons it is desired that confusion be avoided between بن and between بن plus بن , respectively.
- 3. Occasionally the character sequences که, زه, سه محاک and g•h in order to differentiate these romanizations from the digraphs kh, zh, sh, and gh, which are used to represent the characters خ, ڈ, ش خ, غ, ش غ, and ځ.
- 4. Hamzah (*) should be romanized as follows:
- a. In word-initial position, it should not be romanized. In other positions, however, it should be romanized by an apostrophe, e.g. $\dot{+}\dot{+}\rightarrow$ juz'.
- b. Yē with hamzah (¿) should be romanized êy, unless it represents the compound (izafah) morpheme, in which case it is romanized according to rule 10 below.
- 5. The division of words utilized in Pashto writing is followed in romanization, except that the elements —ābād, -khwā, -shahr, -zay and ullāh are always romanized as part of the preceding word, e.g. رَحْمُت الله Raḥmatābād and رَحْمَت الله Raḥmatābād and صاباء Raḥmatābād and and endings of Turkish origin such as —lar, -lī, -lū, -i, -u, -si, and —su, should be written together with the preceding word.

6. The Pashto preposition \ni should be romanized dê in agreement with its pronunciation, despite the fact that it is sometimes pointed with $z\bar{\imath}r$.
7. The word شّ, meaning <god>, should always be romanized Allāh, except as specified in note 5.</god>
8. In names of Arabic origin, the l of the definite article al is assimilated before the 'sun letters' t, \$\bar{s}\$, d, \$\bar{z}\$, r, z, \$\bar{s}\$, \$\bar{s}\$, \$\bar{z}\$, \$\bar{t}\$, \$\bar{z}\$, and n. In romanization, the article should be separated from the name it precedes and should not be capitalized except at the beginning of a name, e.g. خَبَلُ السَّراح Jabal as Sarāj.
9. In names of Arabic origin, a consonant character which would ordinarily be written with a shaddah () in Arabic should be romanized by writing the equivalent consonant letter twice, e.g., مُحَمَّد Миḩammad.
10. In Pashto names which contain the Persian compound (izafah) morpheme, the izafah should be Romanized as –i if the preceding word ends with a consonant other than yeh or silent heh, and as –yi otherwise.
11. The character sequence غ, when followed by ا or عن or when pointed by the vowel character , should be Romanized khw, although the w is either not pronounced, or only weakly pronounced.